stics Statistique
ada Canada

Ottawa

GRAPHICAL RATIONAL PATTERNS (GRP) APPLICATION TO CANADIAN GEOSTATISTICAL DATA

by

J. Yam

E P R I N T E D P U B L I C A T I O N No. 27

inted from: The Canadian Journal of Statistics,
1(2) 1973, 231-246

Canada

Statistics Canada | Statistique Canada

Ottawa

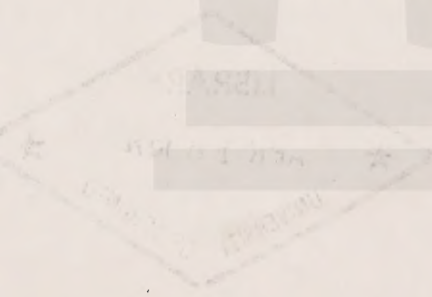
GRAPHICAL RATIONAL PATTERNS (GRP) APPLICATION TO CANADIAN GEOSTATISTICAL DATA

by

J. Yam

REPRINTED PUBLICATION No. 27

Reprinted from: The Canadian Journal of Statistics, 1(2) 1973, 231-246



The Canadian Journal of Statistics
Section C: Applications
Vol.1, No.2, 1973, pp. 231-246

La Revue Canadienne de Statistique
Section C: Applications

GRAPHICAL RATIONAL PATTERNS (GRP) APPLICATION TO CANADIAN GEOSTATISTICAL DATA

by

J.YAM
Statistics Canada

Key Words and Phrases: *Graphical rational pattern, basic characteristics, construction, major symbols, GRP scale, density maps, population distribution maps, internal migration, average intensity, intensity series, base population, trend maps.*

ABSTRACT

This article deals with the application of a new graphical system, known as Graphical Rational Patterns (GRP), to geostatistical data. The first portion of the paper provides a brief review of the system for readers who are unfamiliar with it. This is followed by a series of annotated illustrations based on recent Canadian data. The paper emphasizes that the GRP system is particularly suitable for producing large numbers of inexpensive drafts or experimental maps by non-cartographers; as such, the system contributes to geostatistical analysis undertaken by demographers, statisticians and others. An attempt is made to substantiate that the GRP system *(i)* introduces improved arrays of symbols (*e.g.*, with regard to homogeneity and "rationality"); *(ii)* provides graphs and maps with improved "readability" and transmittance of general trends; *(iii)* suggests solutions where conventional methods are inapplicable or ineffective.

1. INTRODUCTION

One hardly needs to convince statisticians that in analysing statistical data "one picture - that is, a chart, graph or map - is worth a thousand words". This applies particularly to geostatistics[1] where the difficulty of conceiving

1. *The term "geostatistics", not yet used universally, refers to the analysis of geographical-statistical distributions.*

trends embodied in statistical tables is compounded by the need to envisage geographical contexts as well. In this connection, conventional[2] cartographic techniques are inadequate for treating several important categories of series, *e.g.*, double entry tables showing cross-classifications[3]. Even more disturbing, in the author's opinion, is the lack of a general tool enabling analysts who are not professional cartographers to produce large numbers of inexpensive and easily prepared graphs for experiments, drafts, interim reports etc.

The system known as Graphical Rational Patterns *(GRP)* provides some answers to these problems as well as to other deficiencies of conventional techniques, to be discussed later. However, though the *GRP* was introduced several years ago by its inventor, R.Bachi(1968), the system has been applied only rarely, the official publications of Israeli statistics being an exception [see for example the summary volumes 42 and 43 of the 1961 census of Israel]; textbooks on cartography, with the exception of the one by Bertin (1967,p.392),seem to ignore the *GRP* completely.

The aim of this paper is to draw the attention of Canadian statisticians and demographers to the potential of the *GRP* in data analysis. To this end, a series of annotated charts and maps illustrating selected features of *GRP* presentations, are shown in Section 2. For the benefit of readers who are not acquainted with the *GRP* system the remainder of this section reviews its basic properties. A detailed discussion may be found in Bachi(1968) and Yam(1972).

Essentially, the *GRP* system consists of a series of graphical schedules suggested in lieu of conventional ones. The basic *GRP* schedule, termed the single *GRP* scale, provides 101 graphical symbols which correspond to the integers 0-100 *(Fig.1)*. A *GRP* unit, that is, a black square of given area, represents the integer 1. Two, three, ... , nine *GRP*

---

2. *By "conventional" or "traditional" we mean the graphical techniques one is likely to encounter in text-books, official publications, etc.*

3. *Stereograms (also known as "stepped statistical surfaces") may be considered an exception but they are quite impractical on the whole. See, for example, Zelinski (1966, p.8).*

units are clustered in distinct configurations to represent the corresponding integers 2 to 9. The particular cluster configurations are selected to assist in identifying the given symbol, relative to others in the array. The integer 10 is represented by one black square, the area of which equals 10 *GRP* units, that is, ten times the area of the symbol corresponding to the integer 1. The nine configurations employed to show the numbers 1 to 9 are now repeated with the larger squares, thus forming the symbols for 10, 20,...,90 (compare the symbols for 5 and 50, as an example: the configuration is identical but the size of each "building block" is different). A combination of the units and tens thus yields a scale from 1 to 99. Zero is represented by a blank *GRP* and the integer 100 by a solid black square, the area of which equals 100 *GRP* units. The symbols described above are subsequently referred to as *GRP's* or "*GRP* frames". The corresponding integers are termed "basic values".

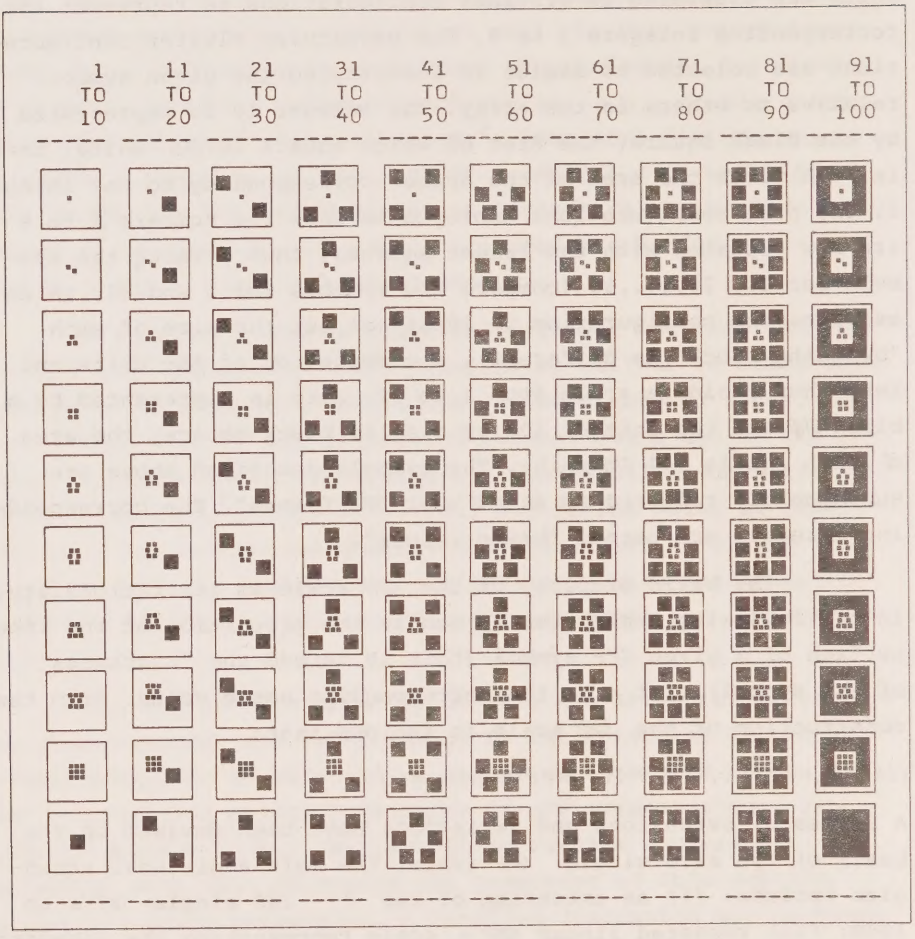
The basic property of the *GRP* scale is its rationality, in the following sense. Let b_n denote the proportion of the inked surface of a given *GRP* symbol (this is termed the "blackness" of the symbol); let n be the corresponding basic value. From the construction of the *GRP* scale it follows that:

(1) $b_0 : b_1 : \ldots : b_n : \ldots = 0 : 1 : \ldots : n : \ldots$

A series of extensions and variations have been devised on the basis of the said single *GRP* scale. The main additional schedules include: *(i)* an extension of the 0 - 100 single *GRP's* to 1000; *(ii)* repeated linear *GRP's*, scale representing the integers 0 to 20 (applied to presentation of traffic lines, for example); *(iii)* repeated areal *GRP's*, scale representing the integers 0 to 10 *(Fig.3)*; *(iv)* multiple *GRP's* for a simultaneous representation of two variables, each in the range 1 to 10.

Additionally, the single *GRP's* may be enclosed in frames of various shapes - square, round, oval, etc.- to introduce an additional qualitative variable or to indicate negative numbers *(Fig.6)*. Various colour schedules increase the versatility of the system further.

FIG. 1 - COMPLETE SCHEDULE OF SINGLE GRP'S, VALUES FROM 1 TO 100



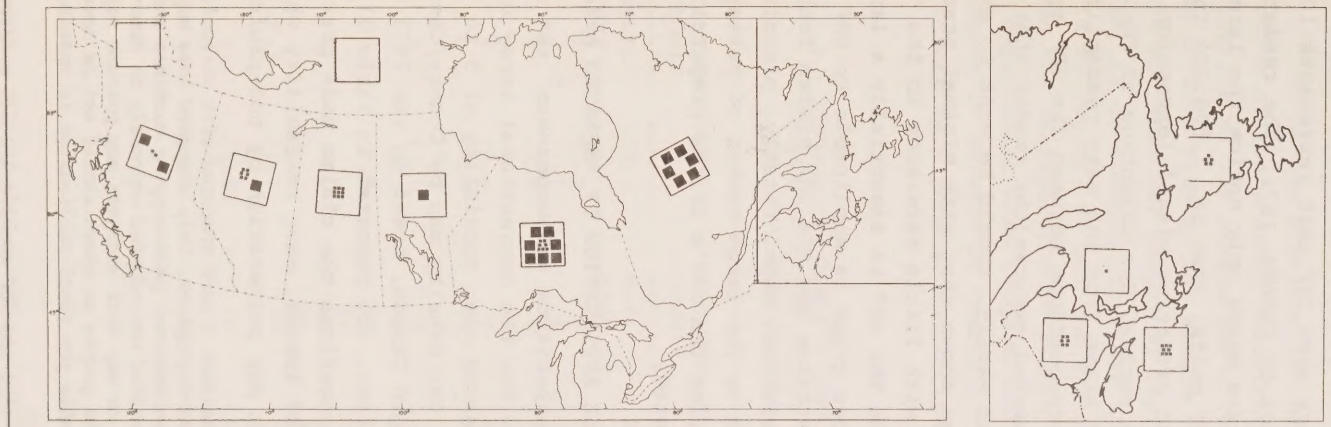
2. ILLUSTRATIONS[4]

2.1. MAPS OF POPULATION DISTRIBUTION - A SIMPLE APPLICATION OF SINGLE GRP'S

According to a common traditional technique, maps showing the regional distribution of a population are prepared with circular symbols, the area of which is proportional to the population in a given region. The *GRP* system proposes to sub-

4. *This footnote is given on page 236.*

FIG. 2 - THE 1971 CENSUS POPULATION OF CANADA BY PROVINCE



KEY: 1 GRP UNIT = 100,000 PERSONS.

EXAMPLES ETC. ETC.

100000 200000 1000000 2000000 7700000 50000 OR LESS

SEE COMPLETE SCHEDULE OF SINGLE GRP'S IN FIG. 1.

NOTE: FOR COMPARISON, THE RAW DATA ARE QUOTED BELOW (IN TERMS OF 1000'S OF PERSONS) -

NFLD. - 522, P.E.I. - 112, N.S. - 789, N.B. - 635, P.Q. - 6,028, ONT. - 7,703, MAN. - 988, SASK. - 926, ALTA. - 1,628, B.C. - 2,185, Y.T. - 18, N.W.T. - 35.

SOURCE: CANADA. STATISTICS CANADA. 1971 CENSUS OF CANADA, POPULATION OF CENSUS DIVISIONS, ADVANCED BULLETIN AP-2. OTTAWA: THE QUEEN'S PRINTER, JUNE, 1972.

stitute single *GRP*'s for the circular symbols. A map of this kind is shown in *Fig.2*, where each *GRP* unit represents 100,000 people in any given province of Canada, 1971. The reader interested in a general picture needs only notice the larger black squares representing 10 *GRP* units, that is, 1,000,000 inhabitants. Thus the *GRP* map conveys an effective visual impression of the major trends latent in the series. One advantage of *Fig.2* over the traditional equivalent stems from the easy readability of *GRP* graphs: the simple correspondence between the *GRP* symbols and the populations represented assures the reader that decoding the symbols may be achieved easily, should he want to exploit this option. Experience has shown that with minimal practice, decoding may be exercised with little reference to the "key" to the map. Furthermore, since the *GRP* is essentially a language, experience acquired from one graph is applicable to *GRP* graphs generally. *Fig 2* also illustrates how the *GRP* system facilitates the preparation of geostatistical maps. Firstly, planning the map is an easy task thanks to the large number of patterns in the array; secondly, applying the *GRP*'s in the preparation proper is a matter of minutes.

2.2. DENSITY MAPS - A SIMPLE APPLICATION OF REPEATED AREAL GRP'S

GRP density maps substitute a homogeneous[5] set of 11 repeated areal *GRP* symbols for the conventional array of patterns - usually a heterogeneous one, consisting of 5 to 7 patterns. *Fig. 3* demonstrates a *GRP* density map by County/Census Division for part of eastern Canada, based on the 1971 Census (the raw data, for comparison, are reported in Table 1). In the author's opinion, the map confirms the claims made previously, concerning transmittance of trends, easy readability when desired, and assistance in map preparation by non-professionals.

---

4. *The illustrations shown in Section 2 were prepared with virtually no assistance of professional cartographers. Only standard base maps, commercially available GRP's, conventional photocopying machines and computer print-outs were used. This method was adopted expressly to demonstrate the type of tentative graph or map that analysts may prepare rapidly in the course of their work. The graphs as shown should not be considered, therefore, as products suitable for final publication in research reports.*

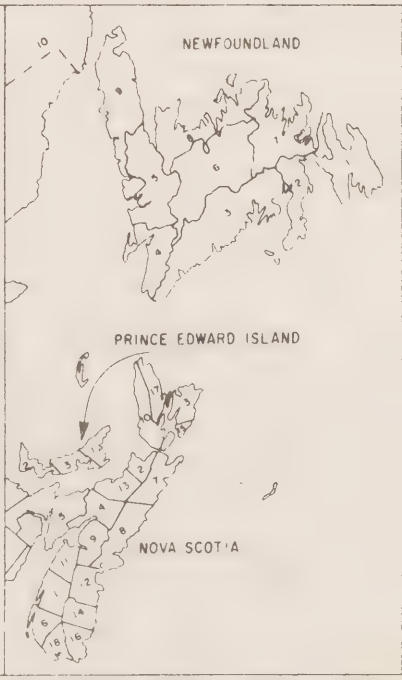
5. *"Homogeneous", in the sense that the GRP symbols are all constructed from the same building blocks, i.e., the GRP units.*

2.3. GRAPHS REPRESENTING DOUBLE ENTRY TABLES

While the previous two examples dealt with *GRP* maps, the objective of the present illustration is to demonstrate a general application of the single *GRP*'s to statistical charting. The virtually uniform size of the single *GRP* frames renders them particularly suitable for graphical representations of double-entry tables. As an example, consider the Canadian internal migration matrix by Region, as per Table 2 and the corresponding *Fig.4*. Though the contribution of graphs to analysis of such data is best appreciated when large matrices are dealt with (*e.g.*, internal migration in Ontario by County), the 5×5 matrix employed here is sufficient as a general indication.

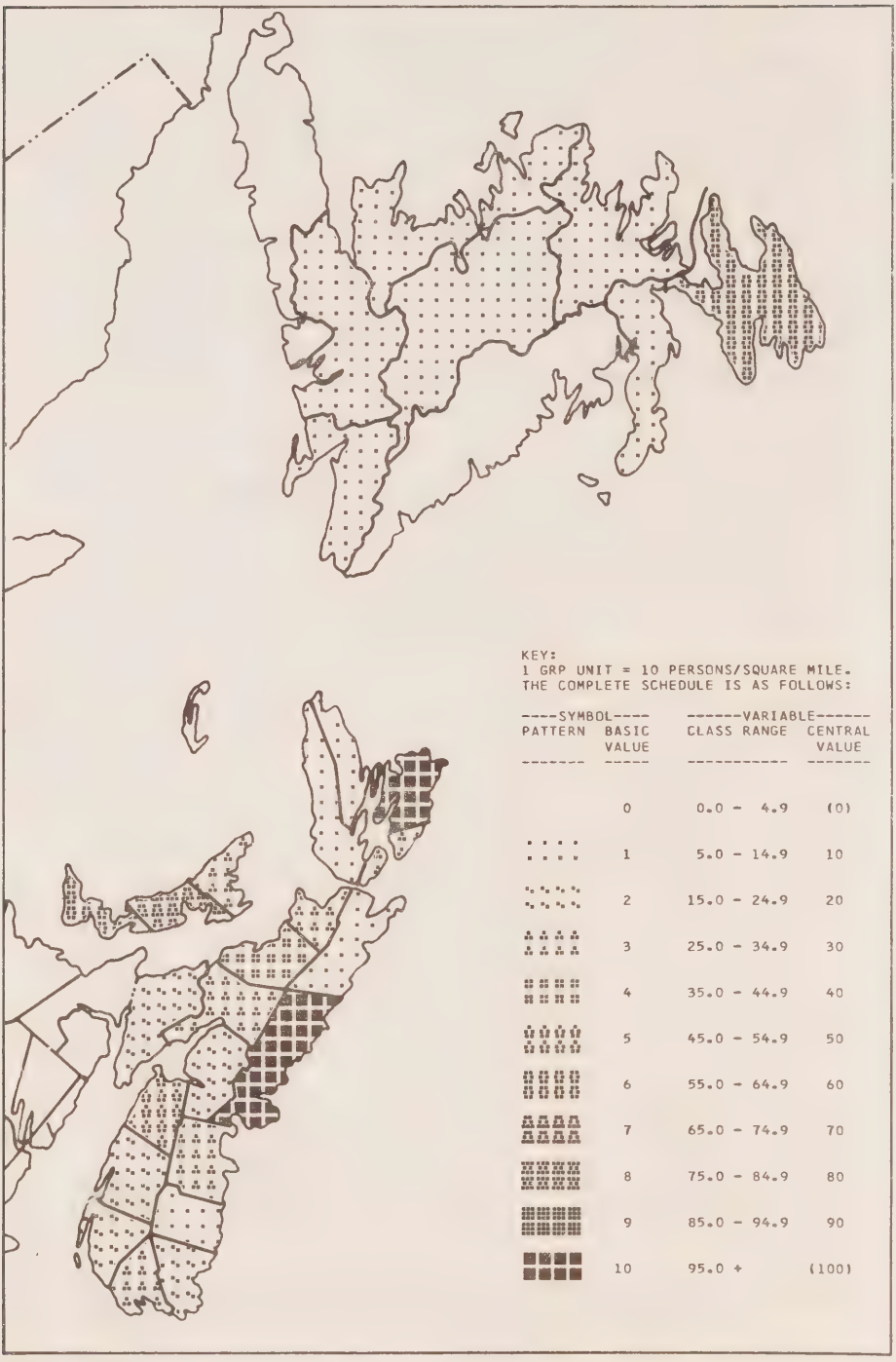
TABLE 1 - POPULATION DENSITY (PERSONS PER SQUARE MILE) BY CENSUS DIVISION (C.D.) OR COUNTY, NEWFOUNDLAND, P.E.I. AND NOVA SCOTIA, 1971

C.D. OR COUNTY	DENSITY	C.D. OR COUNTY	DENSITY
NFLD.	3.6	N.S.	38.7
1	60.6	1	17.7
2	11.1	2	29.8
3	3.0	3	135.2
4	9.8	4	27.0
5	12.6	5	21.2
6	5.7	6	21.3
7	10.5	7	7.6
8	13.6	8	121.8
9	4.1	9	24.5
10	0.3	10	14.3
		11	53.4
		12	34.6
		13	43.1
P.E.I.	51.1	14	14.2
		15	26.8
1	28.5	16	18.3
2	55.1	17	7.3
3	66.0	18	30.9



SOURCE - STATISTICS CANADA, CENSUS DIVISION, GEOGRAPHY SECTION (UNPUBLISHED DATA).

FIG. 3 - POPULATION DENSITY (PERSONS PER SQUARE MILE) BY COUNTY OR CENSUS DIVISION, NEWFOUNDLAND (EXCLUDING LABRADOR), P.E.I. AND NOVA SCOTIA, 1971



SOURCE: TABLE 1.

Examining an internal migration matrix, analysts would typically transform the raw data (Table 2A) into percentages (Table 2B) and proceed to compute "expected" values on the assumption of "independence of marginal totals" (Table 2C). If m_{ij} and e_{ij} denote the "observed" and "expected" streams from i to j (in percentage form), respectively, then:

$$(2) \quad e_{ij} = (\sum_i m_{ij})(\sum_j m_{ij}) /100.$$

The next step may involve the computation of the ratios: "observed"/"expected" (that is, m_{ij}/e_{ij}), to indicate the degree of preference of various streams (Table 2D). Usually, the extreme cases (that is, the most preferred streams as well as the least preferred ones) attract the analyst's attention most.

Figs.4A to *4C* illustrate how these data may be presented in graphical form by using the single *GRP* scales. Referring to the main diagonal of *Fig.4A*, for example, it is readily observed that migration within Regions (according to the definition of Table 2) accounts for the lion's share of internal migration in Canada; from the diagonal of *Fig.4C* it is learnt further that the "observed" streams are at least 2.5 times as large as the "expected" ones for all intra-regional migrations. This example suggests that *GRP* charts may assist the analyst in visualizing the trends latent in complex statistical tables. As mentioned in the introduction, no practical alternative is proposed by conventional cartography.

2.4. MAPS OF SERIES SHOWING "REGIONAL AVERAGE INTENSITY RELATIVE TO A BASE POPULATION" ("INTENSITY SERIES")

Let P_i denote the population of region i as enumerated on one census date; let P_i' correspond to the population at a later census date, so that the percentage change Δ_i is defined by:

$$(3) \quad \Delta_i = 100\ (P_i' - P_i)/P_i \ .$$

For graphical presentations of typical "intensity series", such as the one formed by the values Δ_i, conventional choropleth maps are commonly proposed. This type of map and one of several *GRP* alternatives are demonstrated in *Figs.*5 and 6, respectively.

TABLE 2 - MATRIX OF INTERNAL MIGRATION, CANADA, 1961 CENSUS DATA

A. OBSERVED STREAMS (ABSOLUTE NUMBERS)

FROM REGION	TO REGION ATLNT.	P.Q.	ONT.	PRAIR.	B.C.	CANADA
ATLNT.	170065.	15189.	38621.	6467.	4899.	235241.
P.Q.	8852.	584030.	53941.	6588.	4500.	657911.
ONT.	21056.	40629.	747720.	35575.	17542.	862522.
PRAIR.	5650.	7269.	40990.	424983.	58535.	537427.
B.C.	3328.	3009.	15154.	31226.	225948.	278665.
CANADA	208951.	650126.	896426.	504839.	311424.	2571766.

B. OBSERVED STREAMS (PERCENTAGES)

FROM REGION	TO REGION ATLNT.	P.Q.	ONT.	PRAIR.	B.C.	CANADA
ATLNT.	6.61	0.59	1.50	0.25	0.19	9.15
P.Q.	0.34	22.71	2.10	0.26	0.17	25.58
ONT.	0.82	1.58	29.07	1.38	0.68	33.54
PRAIR.	0.22	0.28	1.59	16.52	2.28	20.90
B.C.	0.13	0.12	0.59	1.21	8.79	10.84
CANADA	8.12	25.28	34.86	19.63	12.11	100.00

C. EXPECTED STREAMS (PERCENTAGES)

FROM REGION	TO REGION ATLNT.	P.Q.	ONT.	PRAIR.	B.C.	CANADA
ATLNT.	0.74	2.31	3.19	1.80	1.11	9.15
P.Q.	2.08	6.47	8.92	5.02	3.10	25.58
ONT.	2.72	8.48	11.69	6.58	4.06	33.54
PRAIR.	1.70	5.28	7.28	4.10	2.53	20.90
B.C.	0.88	2.74	3.78	2.13	1.31	10.84
CANADA	8.12	25.28	34.86	19.63	12.11	100.00

D. RATIOS: OBSERVED / EXPECTED (PERCENTAGES)

FROM REGION	TO REGION ATLNT.	P.Q.	ONT.	PRAIR.	B.C.	CANADA
ATLNT.	889.79	25.54	47.10	14.00	17.20	100.00
P.Q.	16.56	351.16	23.52	5.10	5.65	100.00
ONT.	30.05	18.63	248.71	21.01	16.80	100.00
PRAIR.	12.94	5.35	21.88	402.84	89.94	100.00
B.C.	14.70	4.27	15.60	57.08	669.58	100.00
CANADA	100.00	100.00	100.00	100.00	100.00	100.00

NOTES - A "MIGRANT" IN THIS TABLE IS A PERSON OF AGE 5 OR OVER, WHOSE MUNICIPALITY OF RESIDENCE IN 1961 WAS DIFFERENT FROM THE ONE IN 1956.

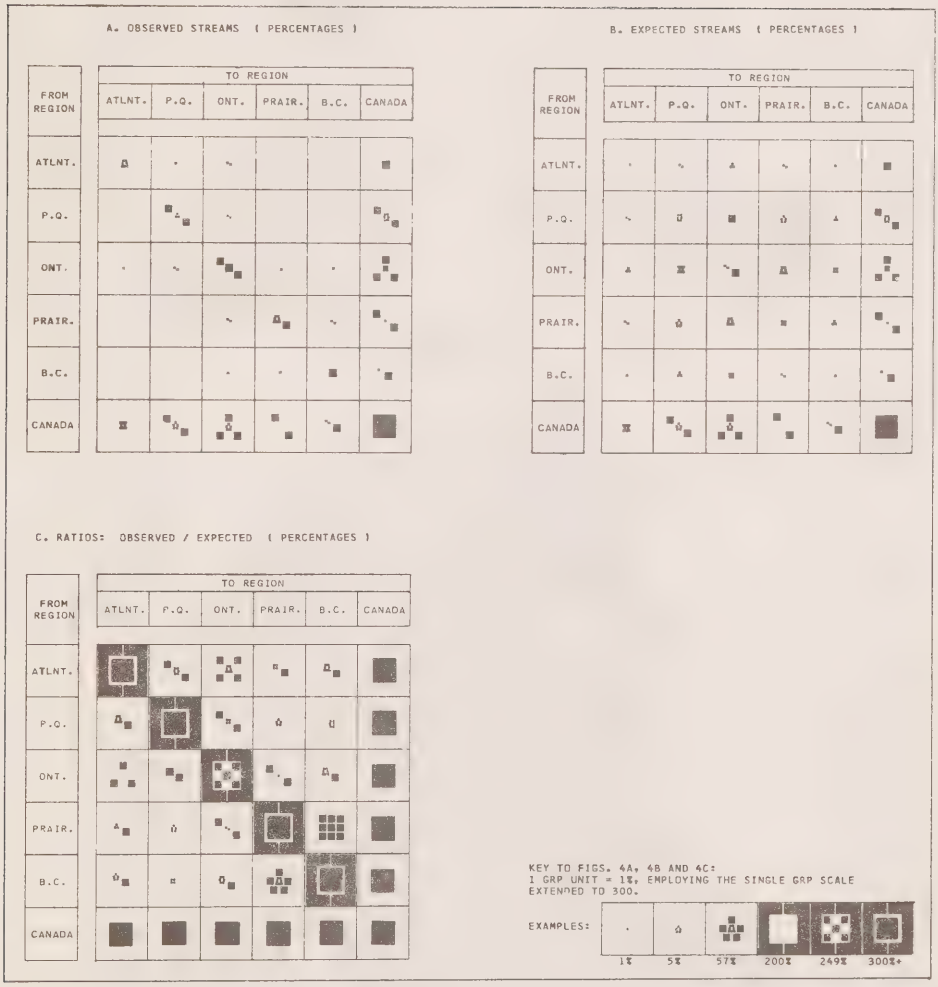
THE TOTAL FOR CANADA IS EXCLUSIVE OF Y.T. AND N.W.T.

TABLE 2 WAS COMPUTED FROM THE FOLLOWING DATA:

(1) FIGURES FOR INTERPROVINCIAL MIGRATION, AS PUBLISHED IN GEORGE, M.V. INTERNAL MIGRATION IN CANADA, TABLE 6.3. OTTAWA: THE QUEEN'S PRINTER, 1970 (A DBS 1961 CENSUS MONOGRAPH).

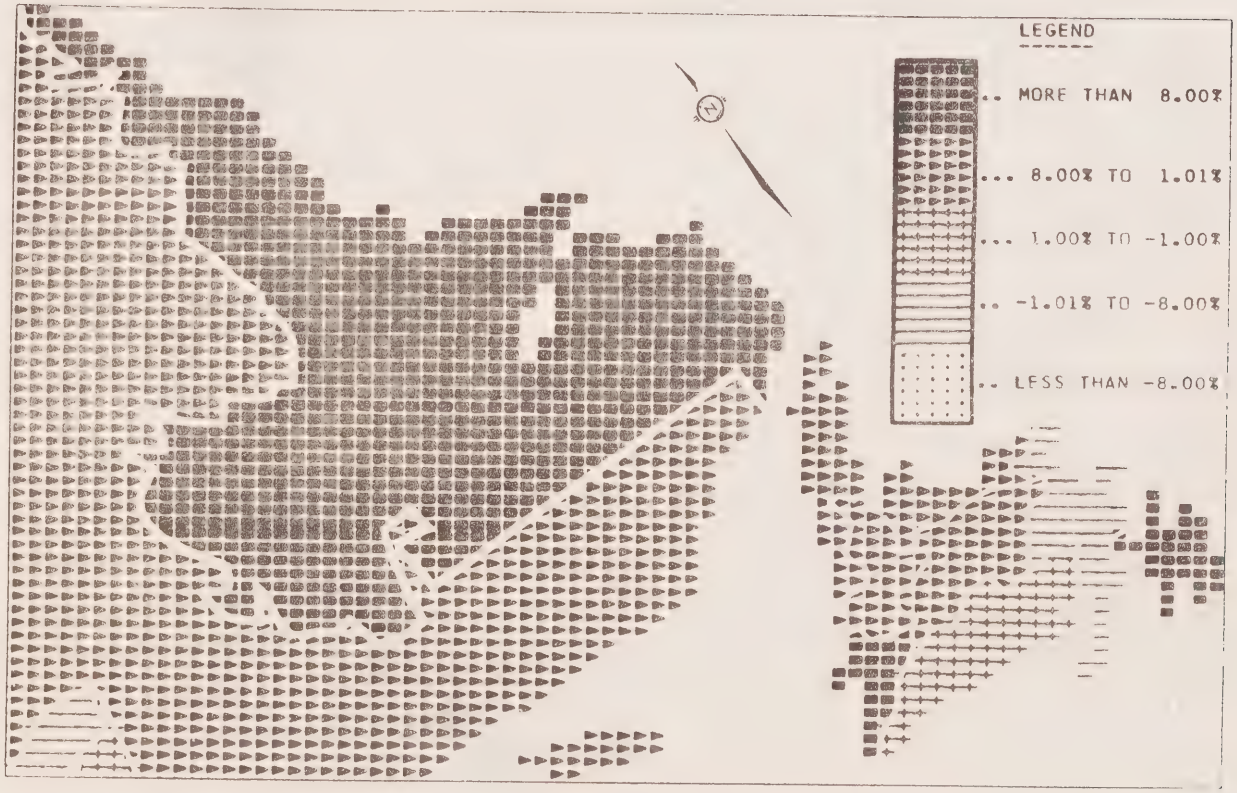
(2) FIGURES FOR INTRA PROVINCIAL MIGRATION, AS PUBLISHED IN CANADA, DBS. 1691 CENSUS OF CANADA, GENERAL CHARACTERISTICS OF MIGRANT AND NON-MIGTANT POPULATION, BULLETIN 4.1-9, TABLE II. OTTAWA: THE QUEEN'S PRINTER, 1965.

FIG. 4 - MATRIX OF INTERNAL MIGRATION, CANADA, 1961 CENSUS DATA



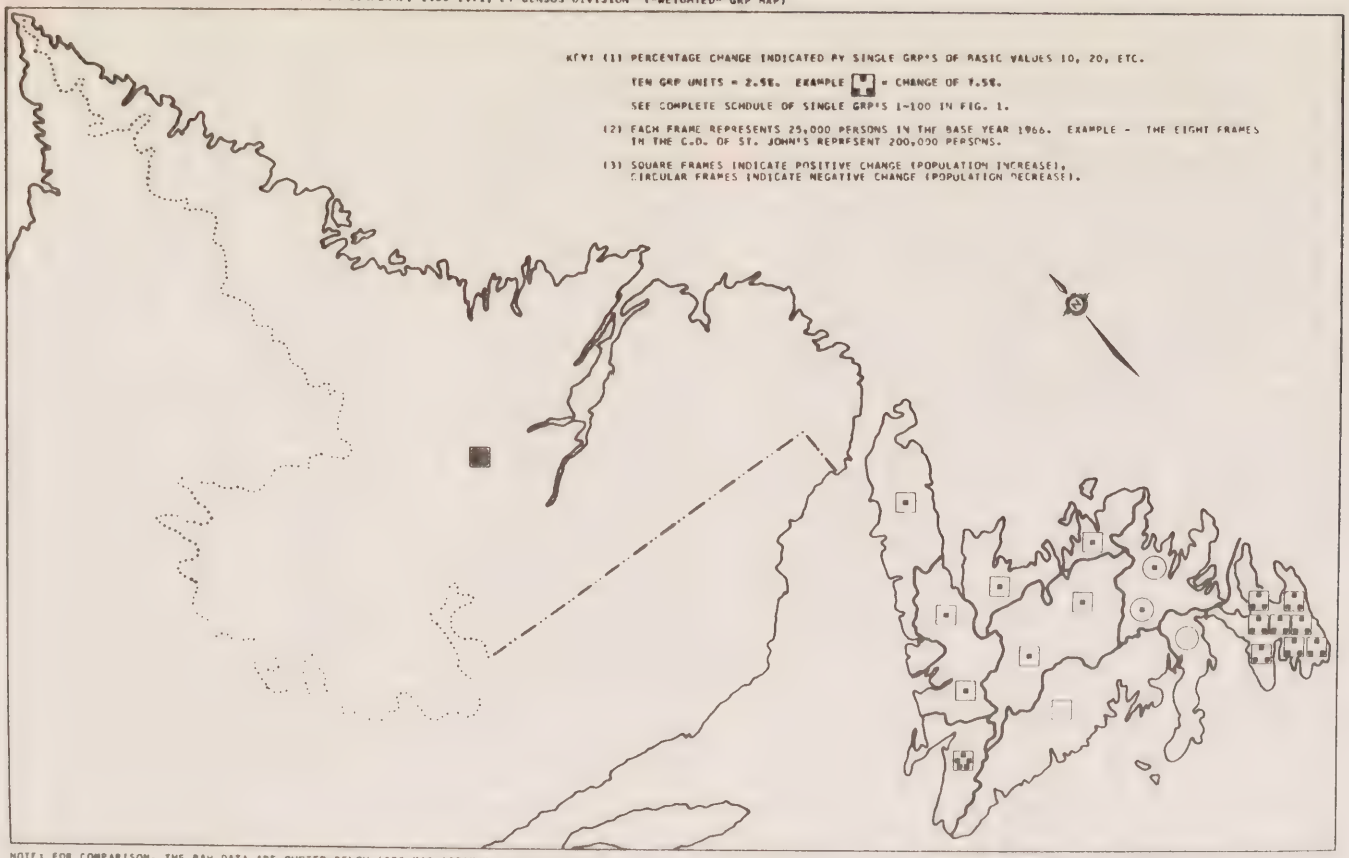
SOURCE: TABLE 2.

FIG. 5 - POPULATION CHANGE (PERCENTAGES) IN NEWFOUNDLAND, 1966 - 1971, BY CENSUS DIVISION
(CONVENTIONAL CHOROPLETH MAP)



SOURCE: ENLARGED SECTION OF THE MAP APPENDED TO THE PUBLICATION CITED AS SOURCE TO FIG. 2.

FIG. 6 - POPULATION CHANGE (PERCENTAGES) IN NEWFOUNDLAND, 1966-1971, BY CENSUS DIVISION ("WEIGHTED" GRP MAP)



NOTE: FOR COMPARISON, THE RAW DATA ARE QUOTED BELOW (SEE MAP ACCOMPANYING TABLE 1) -

SOURCE: AS FOR FIG. 2.

CENSUS DIVISION	NFLD.	1	2	3	4	5	6	7	8	9	10
1966 POP. (1000'S)	493.4	198.3	27.6	24.4	25.3	43.6	39.4	41.4	49.8	22.6	21.2
% CHANGE, 1966-71	+5.8	+8.1	-1.1	+0.5	+12.2	+3.0	+1.9	-1.9	+1.8	+2.5	+33.1

The major difference between the two maps is the visual impression generated by Labrador, and the question arises as to which of the maps transmits correctly the message latent in the data. Clearly, both maps attempt to represent the variable Δ by means of the basic property of the graphical symbols, that is, by their "blackness" (as defined in Section 1). If b_i denotes the blackness of the symbol representing Δ_i, it is reasonable to require that the visual impression generated by region i depend on b_i. In the conventional choropleth map, however, the visual impression is created by weighting the b_i's by the map-area a_i of region i. It is suggested here that *(i)* the weighting factors a_i are irrelevant to the presentation of Δ_i and hence the conventional practice is unjustified; *(ii)* the overwhelming visual impression associated with regions such as Labrador *(Fig.5)* is only in part due to the variable Δ and is mainly a consequence of the region's area - hence the conventional practice may mislead the reader. The simplest *GRP* alternative applies one frame of the single *GRP* schedule per region, such that its b_i is proportional to Δ_i. No weighting factor is introduced and hence the area of the region does not affect the visual impression. An alternative *GRP* technique suggests the preparation of weighted maps for intensity series. Accordingly, p_i frames, rathen than a single one, are assigned to region i: unlike the choropleth maps, however, the weight p_i is determined so that $p_i \propto P_i$. As a result, and recalling expression *(3)*, the total blackness in region i is given by

(4) $$p_i b_i \propto P_i \Delta_i \propto P_i' - P_i .$$

Expression *(4)* implies that the overall visual impression generated in region i is proportional to the magnitude of the intercensal population change (in absolute numbers).

Reviewing *Fig.6*, which represents a weighted *GRP* map, one may easily determine that the value of Δ is higher in Labrdor than in any other Census Division, yet the visual impression created by this region is not as overpowering as in *Fig.5*. On the other hand, and in contrast to *Fig.5*, the map leaves little doubt that the absolute change in Labrador is smaller than the one which took place in the Census Division of St.John's. Summarizing the foregoing discussion it is suggested that *GRP* maps

for intensity series are preferable to conventional presentations.

3. CONCLUSION

In the course of commenting on the four *GRP* maps presented above, an attempt was made to substantiate the propositions that the *GRP* system *(i)* answers the basic requirement of any graphical system, namely, that the essential properties of the data represented be conveyed visually in an effective manner; *(ii)* offers an additional option of easy readability; *(iii)* provides new solutions in certain cases where conventional methods are impractical or otherwise deficient; *(iv)* enjoys additional advantages, theoretical (*e.g.*, rationality) as well as aesthetic (*e.g.*, homogeneity). Significantly, the practicality of the *GRP* as a system which may be applied by non-professionals was emphasized. It seems appropriate, then, to conclude on a practical note, suggesting that statisticians and demographers experiment with this new system in the course of their work.

RÉSUMÉ

Nous considérons l'application d'un nouveau système graphique, connu sous le nom de Graphical Rational Patterns *(GRP)*, à des données géostatistiques. L'article fournit d'abord une brève révision du système et de ses principes. Ceci est suivi par une série d'illustrations cartographiques et par des commentaires. Toutes les cartes sauf une sont basées sur des données canadiennes (principalement du recensement de 1971), l'exception se rapportant aux Etats-Unis. Pour permettre au lecteur de se rendre compte des avantages du système *GRP*, des tableaux des données originales accompagnent chaque carte, et dans trois cas les cartes conventionnelles équivalentes sont ajoutées. L'article accentue le fait que le système *GRP* est particulièrement adapté à la production d'un grand nombre de dessins peu couteux ou de cartes expérimentales par des personnes qui ne sont pas des cartographes; ainsi ce système contribue à l'analyse géostatistique entreprise par des démographes, des statisticiens, etc. D'autres caractéristiques du *GRP* comprennent: *(i)* l'introduction d'un ensemble amélioré de symboles ayant trait à l'homogénéite, "la rationalité", et le nombre de configurations; *(ii)* la préparation de cartes avec une lisibilité accrue et la mise en

évidence des tendances; *(iii)* des possibilités de solutions où les méthodes conventionnelles sont inapplicables ou inefficaces.

REFERENCES

[1] Bachi, R.(1968). *Graphical Rational Patterns.* Israel Universities Press, Jerusalem.

[2] Bertin, J.(1967). *Seminologie Graphique.* Mouton, Paris.

[3] Yam, J.(1972). *The Graphical Rational Patterns System Applied to Geostatistical Maps.* Report submitted to the Study Group on Thematic Maps, Census Division, Statistics Canada.

[4] Zelinski, Z.(1966). *A Prologue to Population Geography.* Prentice Hall, New Jersey.

Received 16 October 1972
Revised 24 January 1973

Dr. J. Norland (Yam)
Demographic Analysis and Research Section
Statistics Canada
Ottawa, Ontario, K1A 0T7
CANADA

CENSUS FIELD PAPERS AND MEMORANDA ISSUED

Reprinted Publications (since 1965)

Census Technical Papers (1965-1968)

Technical Memoranda (1965-1968)

Analytical and Technical Memoranda (since 1969)

Lists of papers published in above series are available from:

Technical Officer
Characteristics Division
Census Field
Statistics Canada
Ottawa, Ontario
K1A 0T7